creative commons

COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

Master's Thesis

Optimal Real-time Monitoring of an Information Source under Communication Costs

Jihyeon Yun

Department of Computer Science and Engineering

Graduate School of UNIST

2019

Optimal Real-time Monitoring of an Information Source under Communication Costs

Jihyeon Yun

Department of Computer Science and Engineering

Graduate School of UNIST

Optimal Real-time Monitoring of an Information Source under Communication Costs

A thesis
submitted to the Graduate School of UNIST
in partial fulfillment of the
requirements for the degree of
Master of Science

Jihyeon Yun

12.28.2018
Approved by

Advisor
Changhee Joo

Optimal Real-time Monitoring of an Information Source under Communication Costs

Jihyeon Yun

This certifies that the thesis of Jihyeon Yun is approved.

12.28.2018

Signature

Advisor: Changhee Joo

Signature

Committee Member: Kyunghan Lee

Signature

Committee Member: Jun Moon

ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY

Abstract

In this work, we study the real-time monitoring of an information source over a communication link with a transmission cost. As the most basic nontrivial setting, we consider a single source governed by a random walk process and design the scheduling policy that minimizes the cost sum of transmission costs and the inaccuracy of the state information at the monitoring center. This problem closely relates to the emerging area of Age-of-Information (AoI) scheduling that focuses on the timely transfer of fresh information over resource-constrained networks. When information updates are controlled by the scheduler, we can incorporate the internal dynamics of the information source into transmission decisions. For a general cost function of estimation errors, we first establish the optimality of a threshold rule for minimizing a weighted sum of the transmission cost and the tracking inaccuracy for the information source. Then, we characterize the optimal threshold level as an explicit function of the transmission cost and source behavior. Further, we provide preliminary results of a learning-based approach that finds the optimal threshold when the transmission cost is unknown a priori.

ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY

ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY

Contents

ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY

List of Figures

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

I Introduction

With the growing number of applications in cyber-physical systems, including autonomous transportation systems, intelligent manufacturing systems, smart power grid applications, etc.,there is an increasing interest and demand for the real-time monitoring of time-varying system status. In such systems, sources send status updates to a destination in order to keep the status state up-to-date. Accordingly, it is desired that updates should be delivered to the destination as fast as possible after they change. The effectiveness of the control decisions at a receiver (destination) significantly depends on the information freshness. However, in reality, the network resources are limited and costly, so a status update from the source might have aged because of the queueing delay and the transmission delay [1]. The information freshness can be measured by an Age-of-Information (AoI) metric, which is defined as how old the recent update is from when it is generated at the source.

Previous works in this domain, for example [1–5], analyzed the age metric to evaluate the performance of the status update system. Age of a status update system was investigated in [1]. It considered a status update system where a transmitter sends status update packets to a receiver through a network cloud. The packets are randomly generated and service times are exponential random variables with some rate. The packets can be delivered in the out-of-order manner, which leads to the complication of computation for status age. The expression of system age was figured out and verified by simulations. Easily expected, its result showed that the system age decreases as we increase the system utilization or the service rate. However, it leads to the costly consumption of network resources. Two sources and one monitor status update system was considered in [2]. Status update packets are delivered through a $M/M/2$ queue and out-of-order packet reception is possible. It analyzed status age metric and compared the result with $M/M/1$ and $M/M/\infty$ queueing system. It has been discovered in these investigations that there is an optimal rate at which a source should transmit its update to a destination to optimize AoI. In [3], the average age for the status update system was analyzed, where there is a single source and channel and system status is updated to a monitor (or monitors) through a first-come-first-served (FCFS) queue. First, age metric under the different FCFS queueing systems ($M/M/1$, $M/D/1$, and $D/M/1$) was formulated. Then, it investigated the optimal utilization, ρ^*, that minimizes average system age, for example, $\rho^* \approx 0.53$ under the $M/M/1$ system, which differs from throughput-maximizing-utilization ($\rho = 1$) or delay-minimizing-utilization ($\rho = 0$). The lower bound for average system age was also provided when the source generates status update packets right after the end of the service for previous packets. The status update system was also considered in [4] in which there are two independent sources and a monitor. Two independent sources share the network resources and they send status updates to the monitor with a FCFS $M/M/1$ queueing strategy. It analyzed the age metric for finding the average status age for the system and proposed the region of feasible status ages for a pair of independent sources. Further, an optimal rate at which the source should generate its updates to the monitor under the existence of interfering load from the other source was provided. Moreover, it was verified in [5] that the channel is frequently idle under the optimal rate. It considered the source which is powered by a stochastic energy harvesting system. Several updating strategies were analyzed using age metric and it was shown that the optimal policy is lazy even when there is sufficient energy for the source to transmit status update.

Based on those analyses, how could we manage the freshness of status updates from the source to the destination using AoI as a metric? In prior works, e.g., [6–9], which investigated this question, a status update was generated as a packet at the source by an exogeneous process. It was the source's responsibility to make decisions on which of the stacked packets to be transmitted to the destination, under the resource limitations. Most prior works, including [6–9], dealt with the question of how to make these decisions to keep the information fresh. It is known that discarding some of the pending packets might be helpful in keeping the freshness of information by bypassing queueing delay for fresher packets at the expense of the loss (or extra delay) of older packets. Status update system where status updates are sampled as a random process and the source can manage the generated samples and decide which packets should be transmitted was considered in [6]. It was shown that average system age will be improved when the source can discard samples, not transmitting to the destination. Not only the average age but also the peak age metric was analyzed in [6]. The multi-class $M/G/1$ queueing system was investigated in [7] in which each sources generate status update packets. It first analyzed peak age-of-information (PAoI) metric for $M/G/1$ and $M/G/1/1$ model. Then, the update rates for both cases was optimized, where the source can choose the update interval to minimize system cost considering PAoI. Minimum age scheduling problem was also handled in [8]. It investigated age of information and proposed the scheduling strategy to optimize the overall information age. To solve the optimality and scalability of its problem, integer linear programming (ILP) formulation and steepest age decent algorithm were employed. Additionally, scheduling problem in wireless network with an aim to minimize age of information and align information at the receiver was investigated in [9]. It analyzed inefficiency of several conventional MaxWeight scheduling policies and developed new scheduling policy combining age with the interarrival times of the incoming packets. Its performance was verified using heavy-traffic analysis. Its new scheduler achieved the state space collapse and the upper and lower bounds for information freshness were suggested.

This expanded model raises the natural question of how to accommodate and combine the information source dynamics into the generation of status updates and transmission decisions. This question has been formulated as a joint optimization problem of scheduling and remote estimation in [10], where the sum of communication cost and average Mean Square Error (MSE) estimation cost is minimized over a finite-time horizon. In [10], a source sends either a real-number or free symbol to an estimator at each time slot. Each real-number transmission incurs a positive cost and free-symbol transmission doesn't require a cost. It was proved that a symmetric threshold policy at the source and the Kalman-filter at the estimator can achieve jointly optimal

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

behavior for the system. The cost minimization of the MSE in remote estimation has been further extended to the case of noise channel with and without communication cost in [11, 12]. The remote estimation problem with a sensor and an estimator was investigated in [11]. Each information transmission requires a cost and the estimator has the estimation error at each time slot. The additive noisy channel was considered so encoding/decoding steps for communication was required. It was shown that a symmetric threshold policy for the sensor and piecewise encoding/decoding policies are optimal to minimize the total cost sum. On the other hand, the multiple communication channels were investigated in [12]. A source has an option to transmit the sensor value using the perfect channel or the additive noise channel, otherwise transmit the free symbol (i.e., no transmission). There are two kinds of constraints which are communication cost constraint and the number of usage constraint. It was shown that threshold-in-threshold policy is the optimal for the setting with a side channel which notices the sign of the underlying state. In an infinite-time horizon, the average MSE cost minimization problem under sampling rate constraint has been addressed in [13], which is close to our work but with a continuous-time process. In [13], a source can update multiple independent wiener process all together or transmit noting at each time. It was verified that event-trigger sampling with a specific threshold can achieve the optimal performance. More recently, the remote estimation problem over a packet-drop channel has been studied in [14]. It is well known that threshold-based policy is the optimal in the remote estimation problem through previous works. Thus, stochastic approximation algorithms to find optimal thresholds were proposed in [14]. Further, the remote estimation problem of the wiener process under the channel delay was considered in [15]. The communication channel has the random delay and the source can manage the sampling of the process under the sampling frequency constraint. It was proved that threshold policy is the optimal for minimum mean square error estimation and found the optimal threshold which is determined by the sampling frequency constraint and the amount of signal variation during the channel delay. These results show that a threshold-based policy can achieve the optimal performance for the MSE.

In this work, we consider an average cost minimization that captures the trade-off between estimation costs and transmission costs. Different from the previous works, we consider the cost minimization problem i) with a more general form of estimation cost rather than the MSE, ii) in the presence of transmission cost, iii) over an infinite-time horizon, and iv) further, our setting is extended to the case when the transmission cost is unknown a priori. We formulate our problem as a discrete-time Markov Decision Process (MDP) and show that a threshold-based policy can minimize the objective function by solving optimal Bellman equations. Given this optimal structure, we then find the optimal threshold value, γ^*, as a function of the transmission cost. We also note that our optimal threshold does not equal that of the policy in [10] due to different characteristics of the information source, which emphasizes the importance of the dynamics of the information source in the control decision.

The remainder of the paper is organized as follows. The system model is described in Section

II. We show in Section III that a threshold-based policy can achieve the optimal performance and find the optimal threshold value γ^*. In Section IV, we evaluate the performance of our threshold-based policy through simulations. Finally, in Section V, we summarize our contributions.

II System Model

Suppose that a transmitter is responsible for communicating a randomly evolving information source to a receiver over a wireless channel. The information is a sensor value that changes according to a random walk process, and can be delivered to the receiver through a wireless transmission. We consider a time-slotted system, where, at each time slot, the transmitter has an option to send the sensor value immediately to the receiver or wait for the next slot. In either case, there is a cost associated with transmission or with stale information at the receiver, and our goal is to minimize the total average cost. We first formulate this problem rigorously in this section.

Let $w(t)$ be an independent and identically distributed (i.i.d.) random process with distribution as

$$w(t) = \begin{cases} 1, & \text{with probability } \theta, \\ 0, & \text{with probability } 1-2\theta, \\ -1, & \text{with probability } \theta, \end{cases} \tag{1}$$

for some $\theta \in (0, \frac{1}{2}]$. Let the sensor state $X_R(t)$ denote a random walk process associated with $w(t)$ as

$$X_R(t+1) := \textstyle\sum_{k=0}^{t} w(k) = X_R(t) + w(t), \quad \text{for} \quad t \geq 0, \tag{2}$$

and $X_R(0) = 0$. At the beginning of time t, the transmitter observes $w(t)$ and can make the transmission decision. In case of transmission, the receiver can update its information to the most recent value, i.e., $X_R(t) + w(t)$. Formally, let $X_E(t)$ denote the (estimated) value at the receiver at the beginning of time t, and let $u(t) \in \{0, 1\}$ be the indicator of transmission decision at time t. If the transmitter transmits the recent information to the receiver at time t (i.e., if $u(t) = 1$), the receiver updates the estimated sensor value, and otherwise, it keeps the old value as

$$X_E(t+1) = \begin{cases} X_R(t+1), & \text{if } u(t) = 1, \\ X_E(t), & \text{if } u(t) = 0. \end{cases}$$

We assume that $X_E(0) = 0$ and there is no transmission error, but each transmission incurs a cost of $c \geq 0$ units. Let $X(t)$ denote the estimation error, i.e., the difference of the sensing value between the transmitter and the receiver at time t, i.e., $X(t) := X_R(t) - X_E(t)$.

We consider the balancing of two types of system cost. One is from the inaccurate sensing value at the receiver and it is assumed to be a monotonically increasing function $f(\cdot)$ of $|X(t)|$ with $f(0) = 0$ and $f(x) \to \infty$ as $x \to \infty$. The other cost is from information updates with per-transmission cost c. For example, if $f(x) = x^2$, we have the MSE with communication cost

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

as in [10]. Our goal is to find a stationary policy that minimizes the total average cost, i.e.,

$$\text{minimize} \quad \lim_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} \left(f\left(|X(t)|\right) + c \cdot u(t) \right). \tag{3}$$

III Design of Optimal Threshold-Based Policy

In this section, we first show that there exists a threshold-based policy that can achieve the optimal performance with cost function $f(\cdot)$, and then find the threshold that minimizes the average cost.

3.1 Optimality of Threshold-Based Policy

We consider a threshold-based policy that makes transmission decisions based on the current value of the information and show that it can achieve the optimal performance. We first model the evolution of $X(t)$ as a discrete-time Markov Decision Process (MDP) with state $|X(t)|$ as shown in Fig. 1. We denote the update decision at state s as action $\alpha_s \in \{0, 1\}$, where $\alpha_s = 1$ implies the information update at state s (i.e., transmission of the sensing value). Let $P(s'|s, \alpha_s)$ denote the transition probability from s to s' under α_s. We also denote the cost under s and α_s as $r(s, \alpha_s) = f(|X(t)|) + c \cdot u(t)$ when action $\alpha_s = u(t)$ is taken at state s at time (or stage) t. In Fig. 1, the tuple of $(\alpha_s, P(s'|s, \alpha_s), r(s, \alpha_s))$ associated with each state transition is shown.

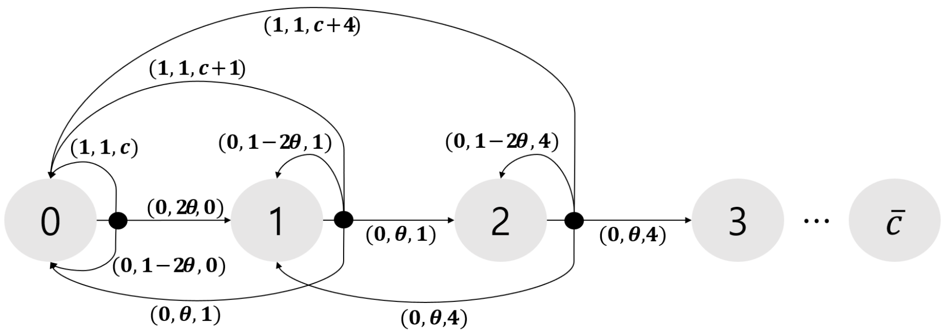


Figure 1: MDP diagram. Each state transition is associated with $(\alpha_s, P(s'|s, \alpha_s), r(s, \alpha_s))$ where $f(x) = x^2$.

We define a stationary policy μ that takes action α_s with probability $\pi(s, \alpha_s)$ at state s. The solution to (3) can be found by minimizing the average cost starting from state 0, i.e., minimizing

$$\lim_{N\to\infty} \tfrac{1}{N} E\left[\textstyle\sum_{t=0}^{N-1} r(s_t, \mu(s_t))\right]. \tag{4}$$

where s_t denotes the state at time t, $\mu(s_t)$ denotes the chosen action at time t, and $s_0 = 0$.

Let μ^* be an optimal policy that minimizes (4). Since $f(s) \to \infty$ as $s \to \infty$, it can be easily found that there exists a constant $\bar{c}$ such that, for all $s \geq \bar{c}$, $\mu^*(s) = 1$, i.e., updating information

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

leads to a lower cost than non-updating. For example, $\bar{c} = \lceil \sqrt{c} + 1 \rceil$ when $f(x) = x^2$. For all $s \geq \lceil \sqrt{c} + 1 \rceil$, the updating action induces the cost of $c + 0^2$. On the other hand, the non-updating action induces the cost of $(s-1)^2$ or $(s+1)^2$ with probability θ or the cost of s^2 with probability $1 - 2\theta$. In either cases, the cost from non-updating is greater or equal to the cost from updating. This implies that the states beyond $\bar{c}$ will never be visited under μ^* if we start from state 0. Thus, we can limit our focus to a finite set of states $\{0, 1, \ldots, \bar{c}\}$.

We consider the optimal Bellman equations to find an optimal policy μ^*. We employ the line of analysis of [16]. Letting each state-transition be a stage and letting $r(s, \alpha_s)$ be the stage cost, our problem is an average cost per stage minimization problem which minimizes (4). From [16], we can solve our problem as an associated stochastic shortest path problem in which there is a cost-free terminal state and visiting the terminal state is inevitable. In stochastic shortest path problem, the terminal state must be visited with minimum expected cost. Since state 0 is visited from any state with non-zero probability, we can set state 0 as the terminal state, at which, upon visit the process terminates. Then, the whole process can be divided into independent cycles marked by visits to state 0. Each of the cycles can be regarded as a state trajectory of a corresponding stochastic shortest path problem with the terminal state. To solve our problem, we first consider a cycle starting from state 0 to the first return to state 0 under a stationary policy μ. Let $N(\mu)$ denote the expected return time, and $Cost(\mu)$ denote the expected cost sum over a cycle under μ. We define λ^* as

$$\lambda^* := \min_\mu \left(\frac{Cost(\mu)}{N(\mu)} \right).$$

Note that λ^* equals the optimal average cost value and thus we have $0 < \lambda^* < c$ (since the policy of updating at every time has the average cost of c). Then, our problem is the same problem that finds a stationary policy μ that minimizes

$$Cost(\mu) - N(\mu)\lambda^* \geq 0. \tag{5}$$

Equation (5) is the expected cost of μ starting from state 0. Minimizing (5) can be viewed as the associated stochastic shortest path problem if we set stage cost to be $r(s, \alpha_s) - \lambda^*$. From [16], we can build bellman equations and there exist optimal costs h^* that solve uniquely the following optimal Bellman equations:

$$h^*(s) = \min_{\alpha_s \in \{0,1\}} \left[r(s, \alpha_s) - \lambda^* + \sum_{s'=0}^{\bar{c}} P(s'|s, \alpha_s) h^*(s') \right],$$

for $s = 0, 1, ..., \bar{c}$. Applying the cost and the state transition probabilities, it can be rewritten as

$$h^*(s) = \begin{cases} \min\left[-\lambda^* + 2\theta h^*(1) + (1-2\theta)h^*(0), \right. \\ \qquad\qquad \left. c - \lambda^* + h^*(0)\right], \quad \text{for } s = 0, \\ \min\left[f(s) - \lambda^* + \theta h^*(s-1) \right. \\ \qquad\qquad +\theta h^*(s+1) + (1-2\theta)h^*(s), \\ \qquad \left. f(s) + c - \lambda^* + h^*(0)\right], \quad \text{for } s = 1, ..., \bar{c}. \end{cases} \tag{6}$$

At each state s, the optimal policy determines α_s that minimizes $h^*(s)$. The following is the main result of this section.

Proposition 1. *There is a threshold-based policy μ^* that minimizes (4), where the policy that chooses action $\mu^*(s_t)$ at time t as*

$$\mu^*(s_t) = \begin{cases} 0, & \textit{for } s_t + w(t) < \gamma, \\ 1, & \textit{for } s_t + w(t) \geq \gamma, \end{cases} \tag{7}$$

for some $\gamma > 0$.

Proof. Let $A(s)$ be the $h^*(s)$ values if we continuously choose optimal action at state s as non-updating ($\alpha_s = 0$). Also, let $A(1) = h^*(1)$ and $A(0) = h^*(0)$. We can express $A(s)$ for $s \geq 2$ as

$$A(s) = f(s) - \lambda^* + \theta A(s-1) + \theta A(s+1) + (1-2\theta)A(s),$$
$$A(s) := \tfrac{f(s)-\lambda^*}{2\theta} + \tfrac{1}{2}A(s-1) + \tfrac{1}{2}A(s+1).$$

Further, let $B(s)$ be the $h^*(s)$ values if we choose optimal action at state s as updating ($\alpha_s = 1$). We have

$$B(s) := f(s) + c - \lambda^* + h^*(0), \quad \text{for } s = 1, ..., \bar{c}.$$

Then, from the Bellman equations, we can set an optimal action as non-updating if $A(s) < B(s)$ and updating if $A(s) \geq B(s)$.

Since we start each cycle from state 0 and non-updating action is a natural decision for small error, we can expect that, during an early period of a cycle, non-updating actions will be chosen and $h^*(s) = A(s)$. We assume that c is not very small, and thus at least at states 0 and 1, non-updating action is optimal. Suppose that in each cycle, non-updating action is continuously chosen until the process arrives at state s. Then Eq. (6) with $\lambda^* > 0$ results in $A(1) > A(0)$ since

$$A(0) = -\lambda^* + 2\theta A(1) + (1-2\theta)A(0),$$
$$2\theta A(0) = 2\theta A(1) - \lambda^*,$$
$$A(0) = A(1) - \tfrac{\lambda^*}{2\theta}.$$

Also,

$$\begin{aligned} A(s+1) - A(s) &= A(s) - A(s-1) - \tfrac{f(s)-\lambda^*}{\theta} \\ &= A(1) - A(0) - \tfrac{1}{\theta}\textstyle\sum_{k=1}^{s}(f(k) - \lambda^*). \end{aligned}$$

From the fact that $f(\cdot)$ is a monotonically increasing function, $A(s+1) - A(s)$ is increasing-then-decreasing or monotonically decreasing (to $-\infty$), and thus, so is $A(s)$. On the other hand, $B(s)$ is a monotonically increasing function. Note that $B(0) = f(0) + (c - \lambda^*) + A(0) > A(0)$ since $\lambda^* < c$. We claim that there is γ such that $A(s) < B(s)$ for all $s < \gamma$ and $A(\gamma) \geq B(\gamma)$. We can prove it by contradiction. If no such γ exists, we should have $A(s) < B(s)$ for all s.

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

According to (6), it implies that the optimal action $\mu^*(s) = 0$ for all $s \in \{0, 1, ..., \bar{c}\}$, which, however, leads to a contradiction since we should have $\mu^*(\bar{c}) = 1$ from our definition of $\bar{c}$. Note that given the existence of such γ, the optimal action changes from non-updating to updating at γ, and once there is an update, a new cycle starts. This implies that function $h^*(s)$ that satisfies the optimal Bellman equations (6) induces a threshold-based policy (7). $\square$

3.2 Closed-Form Optimal-Threshold under Known Costs

In this subsection, we further develop our optimal threshold-based solution by finding the optimal threshold. We can implement our γ-threshold policy given in (7) as follows. At the beginning of each time t, we observe $w(t)$, and calculate $\hat{X}(t+1) := X(t) + w(t)$. If $|\hat{X}(t+1)| \geq \gamma$, the transmitter sends the information (i.e., $u(t) = 1$), and the receiver updates the estimation value (i.e., $X(t+1) = 0$). Otherwise, there is neither transmission nor update (i.e.,$u(t) = 0$) and $X(t+1) = \hat{X}(t+1)$.

Under our policy, process $X(t)$ can be considered as a random walk process that starts at state 0, and when it hits either boundary γ or $-\gamma$, it returns to state 0 as shown in Fig. 2. The expected time for a random walk process to hit the (symmetric) boundary equals the expected length of a cycle in our scenario. Let T denote the time between two consecutive updates. The expected time $E[T]$ can be obtained from [17]. Let $v(x) = E^x[T]$ be the expected time duration to reach state 0 or A when we start at state x in the random walk process like in (1) as in Fig. 2. Then, $v(0) = v(A) = 0$. We have

$$v(x) = \theta\left(1 + v(x-1)\right) + \theta\left(1 + v(x+1)\right) + (1-2\theta)\left(1 + v(x)\right),$$

$$2\theta v(x) = \theta v(x-1) + \theta v(x+1) + 1,$$

$$v(x) = \tfrac{1}{2}v(x-1) + \tfrac{1}{2}v(x+1) + \tfrac{1}{2\theta},$$

$$2v(x) = v(x-1) + v(x+1) + \tfrac{1}{\theta},$$

$$v(x) - v(x-1) = v(x+1) - v(x) + \tfrac{1}{\theta}.$$

Let $d(x) := v(x) - v(x-1)$. Then, we have

$$d(x) = d(x+1) + \tfrac{1}{\theta}.$$

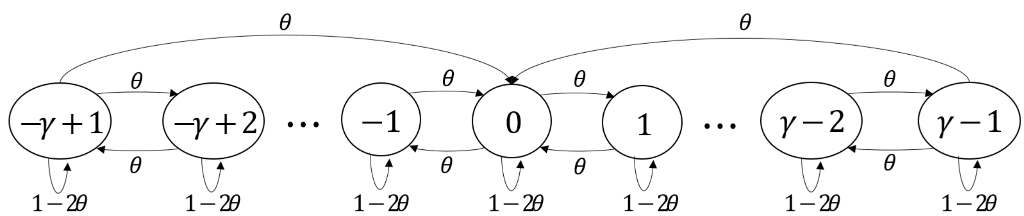


Figure 2: Markov Chain process with states $X(t)$.

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

To solve the equations, we change these equations in a matrix form.

$$\begin{pmatrix} d(x+1) \\ -\frac{1}{\theta} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} d(x) \\ -\frac{1}{\theta} \end{pmatrix}$$

$$\begin{pmatrix} d(x+2) \\ -\frac{1}{\theta} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} d(x+1) \\ -\frac{1}{\theta} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^2 \begin{pmatrix} d(x) \\ -\frac{1}{\theta} \end{pmatrix}$$

$$\begin{pmatrix} d(m) \\ -\frac{1}{\theta} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{m-1} \begin{pmatrix} d(1) \\ -\frac{1}{\theta} \end{pmatrix} = \begin{pmatrix} 1 & m-1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} d(1) \\ -\frac{1}{\theta} \end{pmatrix}$$

From $d(m) = d(1) - \frac{1}{\theta}(m-1)$ and $d(m) = v(m) - v(m-1)$, we have

$$d(m) + d(m-1) + \cdots + d(1) = v(m) - v(0) = v(m).$$

Also, $v(m)$ can be calculated as

$$\begin{aligned} v(m) &= \textstyle\sum_{i=1}^{m} d(i) = \sum_{i=1}^{m} \left(d(1) - \frac{1}{\theta}(i-1)\right) \\ &= m \cdot d(1) - \tfrac{1}{\theta}\left(\tfrac{m(m+1)}{2} - m\right) \\ &= m\left(v(1) - v(0)\right) - \tfrac{m(m+1)}{2\theta} + \tfrac{m}{\theta} \\ &= m \cdot v(1) - \tfrac{m(m+1)}{2\theta} + \tfrac{m}{\theta}. \end{aligned}$$

Using the initial condition, $v(A) = 0$,

$$v(A) = A \cdot v(1) - \tfrac{A(A+1)}{2\theta} + \tfrac{A}{\theta} = 0,$$

$$v(1) = \tfrac{A+1}{2\theta} - \tfrac{1}{\theta}.$$

Using the value of $v(1)$, we have

$$\begin{aligned} v(m) &= m\left(\tfrac{A+1}{2\theta} - \tfrac{1}{\theta}\right) - \tfrac{m(m+1)}{2\theta} + \tfrac{m}{\theta} \\ &= \tfrac{mA - m^2}{2\theta} = \tfrac{m(A-m)}{2\theta}. \end{aligned}$$

Consequentially, we can get

$$v(x) = E^x[T] = \tfrac{x(A-x)}{2\theta}.$$

Remind that $v(x) = \frac{x(A-x)}{2\theta}$ when we start at state x and reach to state 0 or A. To apply this equation to our problem where we start at state 0 and reach to state $-\gamma$ or γ, we can regard our problem as a problem at which we start at state γ and reach to state 0 or 2γ. The expected time for our problem $E[T]$ can be obtained as

$$E[T] = \tfrac{\gamma(2\gamma - \gamma)}{2\theta} = \tfrac{\gamma^2}{2\theta}.$$

Thus, since an update occurs every $\frac{\gamma^2}{2\theta}$ time slots, the average update cost is given by $\frac{2c\theta}{\gamma^2}$.

Next, we estimate the average cost from the information mismatch as a function of γ. Different from the MDP of Fig. 1, we are interested in the distribution of the cost values. We consider a Markov Chain for the evolution of $X(t)$ (i.e., instead of $|X(t)|$) under the γ-threshold rule as shown in Fig. 2. The stationary distribution $\{\pi_i^\gamma\}_i$ of the Markov Chain can be calculated explicitly using global balance equations.

$$\pi_0 = \theta\pi_{\gamma-1} + \theta\pi_{-\gamma+1} + \theta\pi_{-1} + \theta\pi_1 + (1-2\theta)\pi_0$$

$$\pi_1 = \theta\pi_0 + \theta\pi_2 + (1-2\theta)\pi_1$$

$$\pi_{-1} = \theta\pi_0 + \theta\pi_{-2} + (1-2\theta)\pi_{-1}$$

$$\vdots$$

$$\pi_{\gamma-2} = \theta\pi_{\gamma-3} + \theta\pi_{\gamma-1} + (1-2\theta)\pi_{\gamma-2}$$

$$\pi_{-\gamma+2} = \theta\pi_{-\gamma+3} + \theta\pi_{-\gamma+1} + (1-2\theta)\pi_{-\gamma+2}$$

$$\pi_{\gamma-1} = \theta\pi_{\gamma-2} + (1-2\theta)\pi_{\gamma-1}$$

$$\pi_{-\gamma+1} = \theta\pi_{-\gamma+2} + (1-2\theta)\pi_{-\gamma+1}$$

They can be arranged as

$$\pi_0 = \pi_{\gamma-1} + \pi_1,$$

$$2\pi_1 = \pi_0 + \pi_2,$$

$$\vdots$$

$$2\pi_{\gamma-2} = \pi_{\gamma-3} + \pi_{\gamma-1},$$

$$2\pi_{\gamma-1} = \pi_{\gamma-2}.$$

We also have

$$\begin{aligned} 1 &= \pi_0 + 2\pi_1 + 2\pi_2 + \cdots + 2\pi_{\gamma-1} \\ &= \gamma\pi_{\gamma-1} + (2\gamma-2)\pi_{\gamma-1} + (2\gamma-4)\pi_{\gamma-1} + \cdots + 6\pi_{\gamma-1} + 4\pi_{\gamma-1} + 2\pi_{\gamma-1} \\ &= \pi_{\gamma-1}\left(\gamma + \textstyle\sum_{i=1}^{\gamma-1} 2i\right) = \gamma^2\pi_{\gamma-1}. \end{aligned}$$

As a result,

$$\pi_{\gamma-1} = \tfrac{1}{\gamma^2},$$

$$\pi_{\gamma-2} = \tfrac{2}{\gamma^2},$$

$$\cdots$$

$$\pi_1 = \tfrac{\gamma-1}{\gamma^2},$$

$$\pi_0 = \pi_{\gamma-1} + \tfrac{\gamma-1}{\gamma^2} = \tfrac{\gamma}{\gamma^2}.$$

The stationary distribution $\{\pi_i^\gamma\}_i$ is

$$\pi_i^\gamma = \tfrac{\gamma-|i|}{\gamma^2}, \quad \text{for } i \in \{-\gamma+1, \cdots, 0, \cdots, \gamma-1\}.$$

We note that a similar result has been observed in a continuous-time counterpart [13]. Hence, the expected mismatch cost is $\sum_{i=-\gamma+1}^{\gamma-1} f(i) \cdot \pi_i^\gamma$. Combining with the update cost, we have the total average cost given by

$$\frac{2c\theta}{\gamma^2} + \sum_{i=-\gamma+1}^{\gamma-1} f(i) \cdot \pi_i^\gamma \;=\; \frac{2}{\gamma^2}\left(c\,\theta + \sum_{i=1}^{\gamma-1} f(i)(\gamma - i)\right). \tag{8}$$

We can approximate the optimal threshold that minimizes (8) by removing the integer constraint. For example, if we have the MSE cost $f(x) = x^2$, the total average cost becomes

$$\frac{2}{\gamma^2}\left(c\,\theta + \sum_{i=1}^{\gamma-1} i^2(\gamma - i)\right) = \frac{2c\theta}{\gamma^2} + \frac{\gamma^2 - 1}{6}. \tag{9}$$

The optimal threshold which minimizes (9) is $\sqrt[4]{12c\theta}$. The optimal integer threshold γ^* will be determined as either $\lfloor \sqrt[4]{12c\theta} \rfloor$ or $\lceil \sqrt[4]{12c\theta} \rceil$ by comparing (9).

We highlight that our results provide a closed form of the threshold given $f(\cdot)$, and do not require any iterative procedure to obtain the threshold, which is a great advantage to stabilize the system quickly. Also, it clarifies the importance of the dynamics of the information source in characterizing the optimal policy. In large, it seems that as the source has a higher variation (i.e., larger θ), the optimal threshold has to be increased from (8). Further investigation in this direction is an interesting open problem.

3.3 Finding the Optimal-Threshold under Unknown Costs

For the transmitter to solve (8), it is necessary that the cost parameter c is known to the transmitter. In practice, however, the transmission cost may not be known a priori (e.g., when the cost comes from energy consumption at the receiver) or it could even change over time (e.g., when the cost is related to wireless channel state). To this end, we develop a preliminary design of a learning-based γ^*-threshold policy that can find the optimal threshold through learning.

We assume that the per-transmission cost c is fixed and unknown (though its range is known). Although it can be further extended to the cases where the per-transmission cost c_t changes across time t with finite support (following an i.i.d. process with mean $c = E[c_t]$), we assumed a fixed unknown cost for ease of exposition in this work.

Let τ_i denote the time at which the i-th update occurs with $\tau_0 = 0$. The i-the update interval equals $\Delta_i := \tau_{i+1} - \tau_i$, and the average cost $\hat{r}_i$ during interval i can be written as

$$\hat{r}_i = \tfrac{1}{\Delta_i}\left(\textstyle\sum_{t=\tau_{i-1}+1}^{\tau_i} f(|X(t)|) + c\right).$$

We develop a learning-based γ^*-threshold policy that selects threshold $\gamma_i \in [0, \bar{\gamma}]$ for each i-th interval, where $\bar{\gamma}$ is set to the possible maximum threshold. For each possible threshold γ, it maintains three internal parameters: $I(\gamma)$ to evaluate the expected cost, $\eta(\gamma)$ to count the number of selection for γ, and $\hat{r}(\gamma)$ to store the empirical average cost for γ. Specifically, it executes the following procedure at each interval $i > \bar{\gamma}$.

ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY

- At the beginning of the i-th update interval:
  1. For each γ, $I(\gamma) \leftarrow \hat{r}(\gamma) - \sqrt{\frac{2\log i}{\eta(\gamma)}}$.
  2. $\gamma_i \leftarrow \mathrm{argmin}_\gamma I(\gamma)$.
- At the end of the i-th update interval with γ_i:
  1. $\eta(\gamma_i) \leftarrow \eta(\gamma_i) + 1$.
  2. $\hat{r}(\gamma_i) \leftarrow \hat{r}(\gamma_i)(1 - \frac{1}{\eta(\gamma_i)}) + \frac{\hat{r}_i}{\eta(\gamma_i)}$,
     where $\hat{r}_i$ is the empirical average cost
     during the i-th update interval.

For the intervals $i \leq \bar{\gamma}$, it sets $\gamma_i = i$ and updates the parameters $\eta(i), \hat{r}(i)$ as the above.

The learning-based γ^*-threshold policy employs the Multi-Armed Bandit (MAB) technique considering each possible threshold value as an arm. We model each update interval as a time unit of learning and consider the MAB reward as the average cost during the interval. In particular, we use the Upper Confidence Bound (UCB) index [18,19], which is known to achieve the asymptotically optimal performance in the MAB problem, and apply it to our problem to balance the exploration-and-exploitation trade-off in finding the best threshold. We verify the performance of our learning-based γ^*-threshold through simulations.

IV Simulations

In this section, we evaluate the performance of our γ^*-threshold policy through simulations. We first compare our analytical optimal threshold value with the result from Dynamic Programming simulation. Next, we compare our performance with those of two heuristic policies described in Section 4.2. We also evaluate our learning-based γ^*-threshold policy.

Overall procedure is the same for all the policies except their update decision. At each time t, the policy at the transmitter makes an update decision whether the transmitter sends recent information to the receiver ($u(t) = 1$) or not ($u(t) = 0$). The cost during time t is calculated as $f(|X(t)|) + c \cdot u(t)$, with $f(|X(t)|) = |X(t)|^2$ unless otherwise specified. Then, the average cost $AvgCost(t)$ can be written as

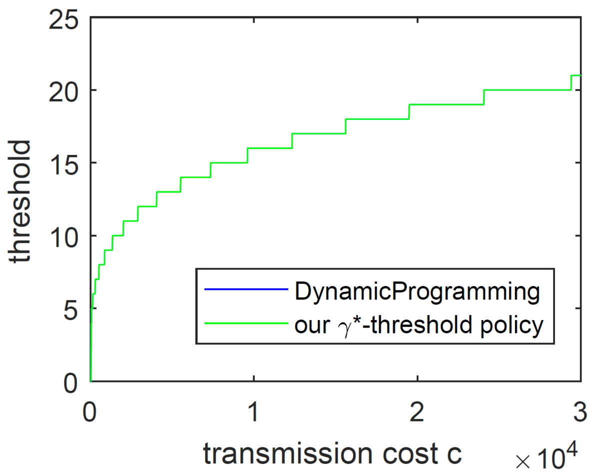
$$AvgCost(t) := \frac{1}{t+1} \sum_{k=0}^{t} \Big(f(|X(k)|) + c \cdot u(k) \Big). \tag{10}$$

4.1 Optimal Threshold

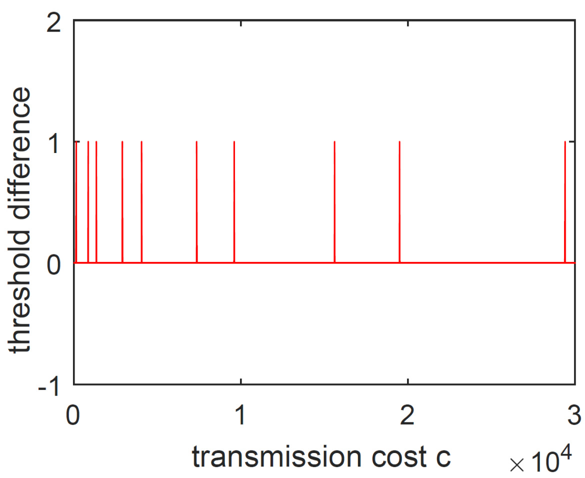
We compare our analytical optimal threshold value with that of Dynamic Programming simulation result. We consider the MSE cost (i.e., $f(x) = x^2$), where our analytical optimal threshold, γ^*, is determined as $\lfloor \sqrt[4]{12c\theta} \rfloor$ or $\lceil \sqrt[4]{12c\theta} \rceil$ by comparing (9). We use value iteration method in Dynamic Programming to find the optimal threshold through simulations [16]. We plot the

UNIST
ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY

thresholds as we change the transmission cost c with fixed source dynamics $\theta = 1/2$. Fig. 3(a) shows the comparison result, at which our analytical optimal threshold is plotted by green line and optimal threshold using Dynamic Programming is plotted by the blue line. Two lines are almost overlapped although there exist few points which are unmatched as shown in Fig. 3(b). It verifies that our γ^*-threshold policy can successfully find the optimal threshold.



(a) Optimal threshold value



(b) Threshold difference

Figure 3: Threshold comparison when $f(x) = x^2$ and $\theta = 1/2$.

4.2 Two Heuristic Policies

For performance comparison, we consider a couple of simple heuristic policies. One is *T-period policy* (or Uniform Sampling in [15]) that updates the information with a fixed period T. Since the performance of T-period policy depends on the value of T, we first find the optimal value T^* as follows.

Note that once the receiver updates the information at time t, $X(t)$ is reset to 0, and thus

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

we can consider $X(t)$ as a renewal process that is reset to 0 with a fixed period. Note that

$$X(T) = \sum_{k=1}^{T} w(k) \ \text{ and } X(0) = 0.$$

From $E[w(t)] = 0$, $E[w(t)^2] = 2\theta$, and the independence of $w(t)$, we have

$$E[X(T)^2] = E[(\sum_{k=1}^{T} w(k))^2] = T \cdot E[w(k)^2] = 2\theta T.$$

Hence, the expected average cost for T-period policy can be obtained by

$$\begin{aligned} E\left[\frac{1}{T}\left(\sum_{k=0}^{T-1} X(k)^2 + c\right)\right] &= \frac{1}{T}\sum_{k=1}^{T-1} E[X(k)^2] + \frac{c}{T} \\ &= \frac{1}{T}\left(2\theta\frac{(T-1)T}{2} + c\right) \\ &= \theta(T-1) + \frac{c}{T}. \end{aligned}$$

$E\left[\frac{1}{T}\left(\sum_{k=0}^{T-1} X(k)^2 + c\right)\right] = \theta(T-1) + \frac{c}{T}$. By minimizing the expected average cost, we can find the optimal time period T^* as

$$T^* = \sqrt{\frac{c}{\theta}}. \tag{11}$$

In our simulations, we denote T^*-period policy as the one that updates the information at every $\sqrt{\frac{c}{\theta}}$ time slots.

We consider another heuristic policy that takes a look one-step ahead in time, denoted by *one-step-ahead policy.* Suppose that the current time is t and the last update decision was at time a (i.e., $X(a+1) = 0$). If we update at time t (i.e., $u(t) = 1$), the average cost of this cycle will be

$$M := \frac{1}{t-a}\left(\left(\sum_{k=a+1}^{t} X(k)^2\right) + c\right).$$

Instead, if we do not update at t and do update at $t+1$ (i.e., $u(t) = 0$ and $u(t+1) = 1$), the average cost of this cycle will be

$$N := \frac{1}{t-a+1}\left(\left(\sum_{k=a+1}^{t+1} X(k)^2\right) + c\right).$$

At each time t, one-step-ahead policy computes M and N, and updates the information if $M < N$. The difference between M and N can be written as

$$\begin{aligned} N - M &= \frac{1}{t-a+1}\left(\left(\sum_{k=a+1}^{t+1} X(k)^2\right) + c\right) - \frac{1}{t-a}\left(\left(\sum_{k=a+1}^{t} X(k)^2\right) + c\right) \\ &= \frac{1}{t-a+1}\left(\left(\sum_{k=a+1}^{t+1} X(k)^2\right) + c\right) - \frac{1}{t-a+1}\left(\left(\sum_{k=a+1}^{t} X(k)^2\right) + c\right)\left(1 + \frac{1}{t-a}\right) \\ &= \frac{X(t+1)^2}{t-a+1} - \frac{\sum_{k=a+1}^{t} X(k)^2}{(t-a+1)(t-a)} - \frac{c}{(t-a+1)(t-a)} \\ &= \frac{1}{t-a+1}\left(X(t+1)^2 - \frac{\left(\sum_{k=a+1}^{t} X(k)^2\right) + c}{t-a}\right). \end{aligned}$$

At the beginning of each time t, the transmitter observes $w(t)$ and has $\hat{X}(t+1) = X(t) + w(t)$, which is the amount of the information mismatch in the case of no information update decision at time t. Thus, the one-step-ahead policy makes the transmission decision $u(t)$ as

$$u(t) = \begin{cases} 1, & \text{if } \hat{X}(t+1)^2 \geq \frac{\sum_{k=a+1}^{t} X(k)^2 + c}{t-a}, \\ 0, & \text{if } \hat{X}(t+1)^2 < \frac{\sum_{k=a+1}^{t} X(k)^2 + c}{t-a}. \end{cases} \tag{12}$$

ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY

4.3 Performance Comparison

We consider a time-slotted system with one transmitter and one receiver, where the transmitter monitors a random walk information source and updates the information at the receiver through wireless transmissions. We assume that there is no transmission error, and both the transmitter and the receiver have the same initial sensing value with $X_R(0) = X_E(0)$ (i.e., $X(0) = 0$).

We compare the performance of the three policies (i.e., γ^*-threshold, T^*-period, and one-step-ahead) in terms of their average costs. We set the threshold of γ^*-threshold policy to the value that minimizes (8), and the period of T^*-period policy as in (11). The decision of one-step-ahead policy is determined dynamically as in (12). We run 100 simulations for each policy, where each runs for 1000 time slots. The results shown in Fig. 4(a) are the average of 100 simulations when $c = 50$ and $\theta = \frac{1}{2}$. We can observe that γ^*-threshold policy achieves the lowest average cost.

Fig. 4(b) shows the results under the same simulation settings except different per-transmission costs c. We show both the simulation results (marked with (sim)) and the analysis results (marked with (cal)) for comparison. The analysis results are presented by a dashed line, and they are well matched with the simulation results. It shows that our γ^*-threshold policy outperforms T^*-period policy and one-step-ahead policy, and the performance differences enlarge as the per-transmission cost increases.

Fig. 4(c) shows the results with different source dynamics θ. We change θ while c is fixed to 50. As expected, the higher variation (i.e., higher θ) the information source has, the larger average cost we have. In all the cases, our γ^*-threshold policy achieves the lowest average costs. Also, for the γ^*-threshold policy, we can observe small jumps in the simulation results around $\theta = 0.03, 0.135$, and 0.43. Those jumps are due to the integer constraint of the threshold, and thus they are not shown in the analysis results. Indeed, similar jumps can be observed in Fig. 4(b) around $c = 13, 42, 104, ...$, which, however, are less noticeable due to the larger scale of the y-axis.

We also evaluate the performance of γ^*-threshold policy with different estimation error function $f(\cdot)$. We set $f(x) = |x|^\alpha$ and change α in $[0.1, 3]$. Fig. 5(a) shows γ^* for different α's when $c = 50$ and $\theta = \frac{1}{2}$, which gradually decreases as α increases. Fig. 5(b) shows the average cost ranges of γ^*-threshold policy for 100 simulations, which are linearly increased with α. We omit the analysis results since they are very similar to the simulation results.

Next, we evaluate our learning-based γ^*-threshold policy with unknown transmission costs. We consider two cases that the transmission cost c is fixed within $[1, 1000]$, and that the transmission cost c_t independently changes across time t. Also, we set $\theta = \frac{1}{2}$ and use the MSE cost function $f(x) = x^2$. The transmitter uses our learning-based γ^*-threshold policy presented in Section 3.3.

Fig. 6(a) shows the performance of our learning-based policy under fixed transmission cost $c = 50$ for $3 \cdot 10^7$ time slots. For the comparison purpose, we also simulate γ^*-threshold policy

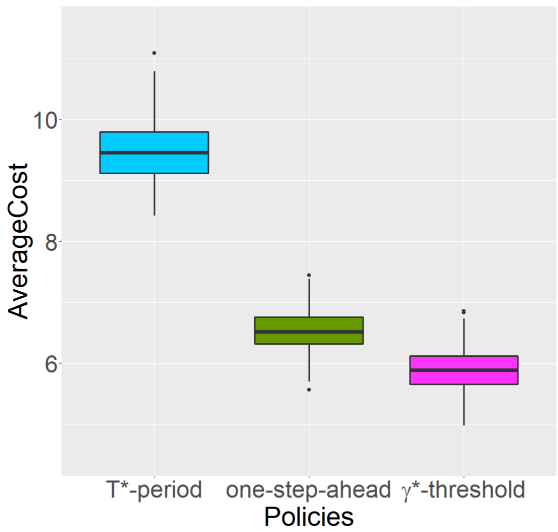
ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY

assuming that c is known and show its results together (dotted line in Fig. 6(a)). It verifies that learning-based policy successfully finds the optimal threshold, and as a result, its average cost converges to that of γ^*-threshold policy that knows the transmission cost c. For the case when transmission cost c_t independently changes across time t, the results are almost the same. We choose c_t following the binomial distribution $B(100, \frac{1}{2})$, with $E[c_t] = 50$. As shown in Fig. 6(b), the results are very similar with Fig. 6(a) and learning-based γ^* policy works well in both cases. Finally, Fig. 6(c) illustrates the changes of the internal parameters $I(\gamma)$ for $\gamma = 3, 4, 5, 6$. The other parameters $I(\gamma)$ with $\gamma \notin [3, 6]$ are chosen rarely and thus omitted. We show them in unit of interval. Fig. 7(a) and 7(b) show that $\gamma = 4$ and $\gamma = 5$ are chosen most frequently. As shown in Fig. 6(c), it also can be shown using the fact that the parameter values are gradually decreasing when γ is not chosen, and abruptly increasing when it is chosen, as expected from its algorithm in Section 3.3.

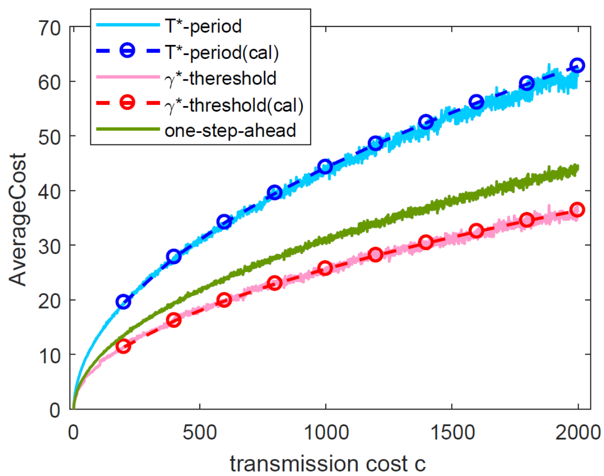
V Conclusion

We investigated the optimal status-update policy where the source sends status updates to the destination under a random-walk-driven evolution model. We focused on the cost trade-off between information freshness and wireless transmission. Under the quadratic-form cost of information accuracy and the per-transmission cost, we showed that a threshold-based policy can minimize the total average cost sum, and characterized the optimal threshold level γ^* based on the information source dynamics and the transmission cost. Through simulations, we compared our γ^*-threshold policy with other two intuitive heuristics and demonstrated that our policy outperforms the others. We also developed a learning-based γ^*-threshold policy to find the optimal threshold γ^* for the cases that the transmission cost c is unknown. We verified that our learning-based γ^*-threshold policy successfully achieved the minimal cost. Theoretic proof for the optimality of the learning-based γ^*-threshold policy remains as a future work. Another interesting research direction is to characterize optimal update policies under different source dynamics.

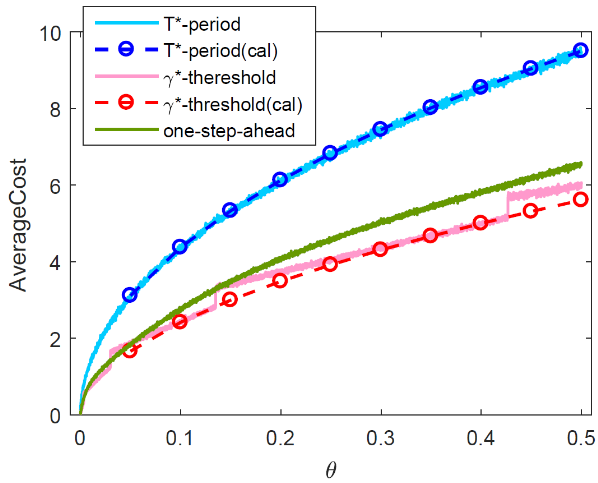
ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY



(a) When $c = 50$ and $\theta = \frac{1}{2}$



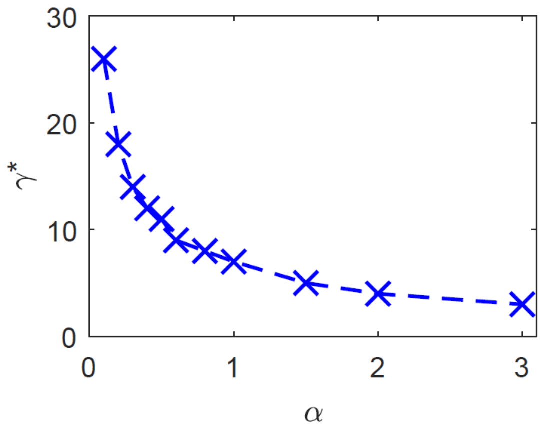
(b) When $\theta = \frac{1}{2}$



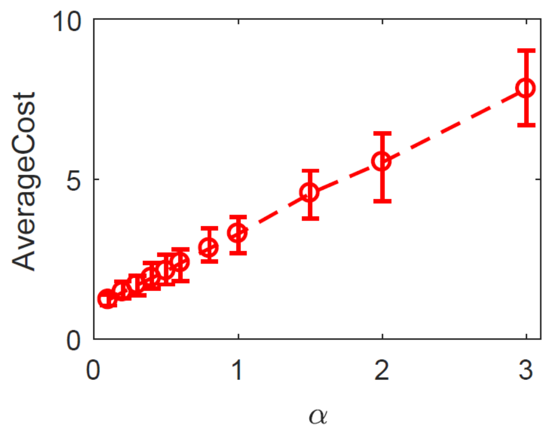
(c) When $c = 50$

Figure 4: Performance of the three policies.

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY



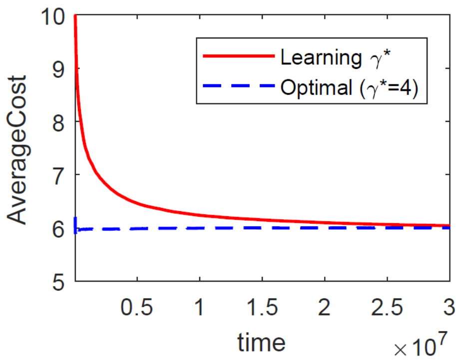
(a) Optimal γ^*



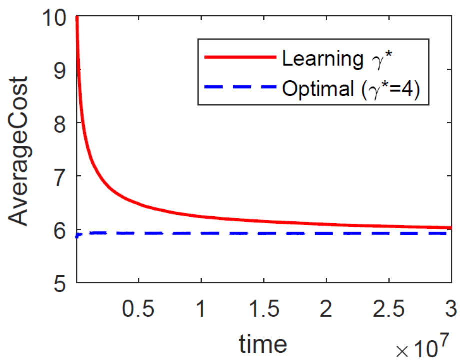
(b) Average costs

Figure 5: Performance of γ^*-threshold policy with $f(x) = x^\alpha$.

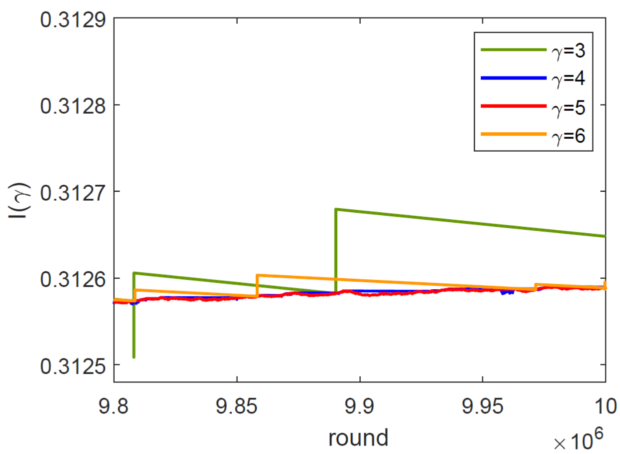
ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY



(a) Average cost with fixed c



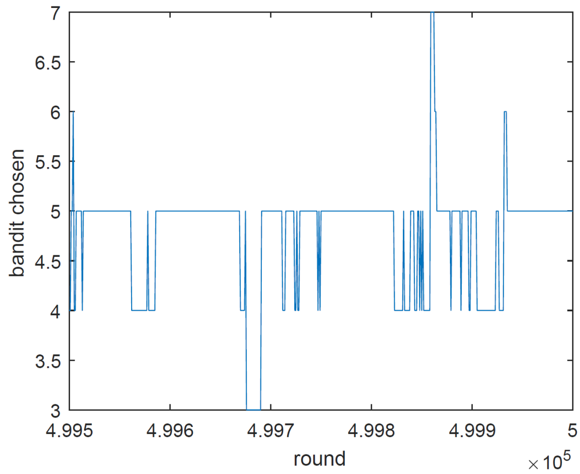
(b) Average costs with varying c_t



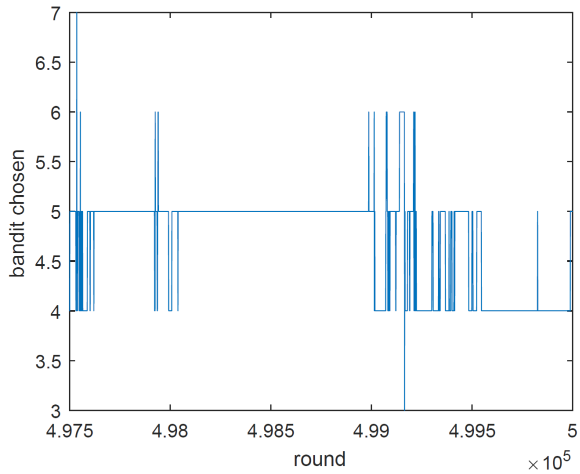
(c) Internal paras. $I(\gamma)$ with fixed c

Figure 6: Performance of the learning-based γ^*-threshold policy.

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY



(a) Bandit choice with fixed c



(b) Bandit choice with varying c_t

Figure 7: Bandit choice of γ^*-threshold policy.

ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY

References

[1] C. Kam, S. Kompella, and A. Ephremides, "Age of information under random updates," in *IEEE ISIT*, 2013.

[2] ——, "Effect of message transmission diversity on status age," in *IEEE ISIT*, 2014.

[3] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *IEEE INFOCOM*, 2012.

[4] R. D. Yates and S. Kaul, "Real-time status updating: Multiple sources," in *IEEE ISIT*, 2012.

[5] R. D. Yates, "Lazy is timely: Status updates by an energy harvesting source," in *IEEE ISIT*, 2015.

[6] M. Costa, M. Codreanu, and A. Ephremides, "Age of information with packet management," in *IEEE ISIT*, 2014.

[7] L. Huang and E. Modiano, "Optimizing age-of-information in a multi-class queueing system," in *IEEE ISIT*, 2015.

[8] Q. He, D. Yuan, and A. Ephremides, "Optimizing freshness of information: On minimum age link scheduling in wireless systems," in *IEEE WiOpt*, 2016.

[9] C. Joo and A. Eryilmaz, "Wireless scheduling for information freshness and synchrony: Drift-based design and heavy-traffic analysis," in *IEEE WiOpt*, May 2017.

[10] G. M. Lipsa and N. C. Martins, "Remote state estimation with communication costs for first-order lti systems," *IEEE Transactions on Automatic Control*, vol. 56, no. 9, pp. 2013–2025, Sept 2011.

[11] X. Gao, E. Akyol, and T. Başar, "Optimal sensor scheduling and remote estimation over an additive noise channel," in *American Control Conference*, July 2015.

[12] ——, "On remote estimation with multiple communication channels," in *American Control Conference (ACC)*, July 2016.

[13] K. Nar and T. Başar, "Sampling multidimensional wiener processes," in *53rd IEEE Conference on Decision and Control*, Dec 2014.

ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY

[14] J. Chakravorty, J. Subramanian, and A. Mahajan, "Stochastic approximation based methods for computing the optimal thresholds in remote-state estimation with packet drops," in *American Control Conference (ACC)*, May 2017.

[15] Y. Sun, Y. Polyanskiy, and E. Uysal-Biyikoglu, "Remote estimation of the wiener process over a channel with random delay," in *IEEE ISIT*, June 2017.

[16] D. P. Bertsekas, *Dynamic programming and optimal control.* Athena scientific Belmont, MA, 1995, vol. 1, no. 2.

[17] O. C. Ibe, *Elements of Random Walk and Diffusion Processes.* John Wiley & Sons, 2013.

[18] P. Auer, N. Cesa-Bianchi, and P. Fischer, "Finite-time analysis of the multiarmed bandit problem," *Mach. Learn.*, vol. 47, no. 2-3, May 2002.

[19] S. Kang and C. Joo, "Low-complexity learning for dynamic spectrum access in multi-user multi-channel networks," in *IEEE INFOCOM*, April 2018.

ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY

Acknowledgement

The research included in this thesis could not be completed successfully if there is no assistance, patience, and support of many individuals.

Foremost, I would like to express my deepest gratitude to my advisor, Prof. Changhee Joo, for the good advice and encouragement which have been continuously given to me from my Master's entrance. He gave me the insight into how to solve difficulties for the research and he has helped me through hard times over the steps for the analysis and the writing of the thesis. I again express my sincere gratitude for that.

I appreciate my thesis committee members, Prof. Kyunghan Lee and Prof. Jun Moon, for giving me instructive advice. I could learn many things from them including lectures that they gave for students.

This research could not have been performed without the cooperation of Prof. Atilla Eryilmaz. I sincerely thank him for the continuous interest in my research even though he is far away.

I would like to thank Jiyeong Hong, who is my valuable friend, for making it through a rough time together. It would be impossible for me to finish my Master's Program without her. She always pays attention to my voice with her warm heart.

I would also like to express my appreciation to Yuri Hyun who is my oldest and dearest friend. We have an emotional connection although the distance between us is far. I'm always thankful for her continued support.

I would like to thank my loving friend, Jisu Kwon. All the time, I feel like we were on the same wavelength. I really appreciate her cheer and encouragement during Master's course.

I would additionally like to thank my colleague, Eunbi Seomoon and Hyein Yu. Two years for Master's degree would be a long and lonely time without them. They always empathized with me, which made me encouraged.

I would also like to thank my seniors, Seonggyu Choe and Sehoon Kim, who brought me laughter whenever I was depressed. They give me strength in times of adversity.

Finally I would like to extend my deepest gratitude to my parents, Byeongcheol Yun and Misun Kim, and my dearest brother, Dokyun Yun. I really love you so much and you are my reason for devoting myself to my studies. I always thank you for your sincere consideration and unconditional love.